The Lagrange Inversion Theorem

Anwar Layada

May 2024

We explore the Lagrange inversion theorem and its applications to formal power series and complex analysis with detailed worked examples. We provide a proof of the theorem in terms of the Cauchy coefficient formula as well as a novel proof of a generalisation of Vandermonde's convolution.

Contents

1 Introduction

The Lagrange inversion formula is a fundamental result that allows us to relate the coefficients of an unknown power series to those of a known power series. In its simplest form, it solves explicitly for the power series coefficients of an unknown $y = y(x)$ given by the implicit equation $y = x\phi(y)$ in terms of the power series coefficients of a known $\phi(z)$. The Lagrange inversion formula, and the more general Lagrange-Bürmann formula, are often stated as results of analysis. Our account will focus on its applications to formal power series (see Section 2.1 for definitions) and consider its applications to analysis as a special case. Our proof of the theorem is based on the proof seen by [4] by Flajolet and Sedgewick with additional exposition, both original and from the proof provided by Erlang Surya and Lutz Warnke [9]. Although this proof relies on statements of complex analysis, we will see that our assumptions hold generally for formal power series. Additional proofs and greater exposition can be found in [6] by Gessel.

In this section, we will informally discuss why solving equations of this form is useful. This will be done more formally in later sections where we will discuss the differences between the analytical and combinatorial interpretations of the formula. We find that many problems find themselves applicable to the Lagrange inversion theorem. The form we will be focusing on is also known as the Lagrange-Bürmann formula.

We will briefly discuss some combinatorial and analytical applications of the Lagrange inversion theorem, with explicit examples coming later. Equations of the form $y = x\phi(y)$ naturally appear in enumerative combinatorics. We will see in Section 4.2, how the Lagrange inversion theorem directly produces counting numbers from generating functions.

Given an analytic function, $F(x)$, Lagrange inversion can provide the compositional inverse. Defining and substituting $\phi(x) := x/F(x)$ in $y = x\phi(y)$ yields $y = xy/F(y)$ or $F(y) = x$. Assuming it exists, this setup allows $y(x)$ to be the compositional inverse of $F(x)$, where a closed-form[2] inverse may not exist. For 'nice' examples, this substitution is equivalent to a simple rearrangement. When working with functions, one must take care when considering the domain and co-domain. In Section 4.1, we produce the Lambert W Function, the compositional inverse of $f(y) = ye^y$, which cannot be expressed in terms of elementary functions[2, 1]. The insolubility of the quintic is a well known result from Galois theory[3, p.65]. Lagrange inversion allows us to produce compositional inverses of a polynomial, $f(x)$. Evaluating these inverses at some a within their radius of convergence provide us with explicit solutions of the polynomial $f(x) = a$ in terms of a convergent infinite sum. Again, one must take care with the domain and co-domain.

In Section 5, we present an original proof of a generalisation of Vandermonde's convolution. It is an excellent showcase of the Lagrange inversion formula's ability to produce non-trivial identities by considering the same equation in trivially different ways (this is also done when we derive Abel's binomial theorem). For this result, we also prove and use an equivalent form of the Lagrange inversion theorem, provided and proved in Section 3.3, that is vital for our proof to work. There are cases when the standard Lagrange inversion formula produces an overly complicated result. There exist many equivalent forms of the Lagrange inversion formula that often allow us to by-pass these complications as we will see later. We state, use, and prove some of these forms, referring the reader to [6, Section 2] for other equivalent forms.

2 Background

The reader is expected to be comfortable working with manipulations of elements of the ring of formal power series. This includes applications of the binomial theorem or reordering and relabeling sums. We will state without proof or exposition results of complex analysis that will later be used in our proof.

2.1 Ring of Formal Power Series

Intuitively, a formal power series can be thought of as a power series where we ignore all notions of convergence and merely focus on the coefficients. In this context, we rarely care for any potential evaluation of the series at a given value. Given a ring R, the ring of formal power series $R[[x]]$ is the set with all elements of the form

$$\sum_{n=0}^{\infty} r_n x^n, \quad r_n \in R$$

with addition and differentiation being defined term wise and multiplication defined via the distributive property, as one would naturally expect.

$$\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n := \sum_{n=0}^{\infty} (a_n + b_n) x^n, \tag{2.1}$$

$$\sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} b_n x^n := \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) x^n, \tag{2.2}$$

$$\left(\sum_{n=0}^{\infty} a_n x^n \right)' := \sum_{n=0}^{\infty} a_n (x^n)' = \sum_{n=1}^{\infty} n a_n x^{n-1}. \tag{2.3}$$

Many properties of the derivative that apply to standard power series, such as the chain rule and product rule, also apply as expected to formal power series.

In future applications, we will find that the coefficients of these series have their own worth with the power series merely being a means of generating them.

2.2 Coefficient Operator

We use $[x^n]f(x)$ to denote the coefficient of x^n within the formal power series $f(x)$.

$$[x^n]f(x) = f_n, \quad \text{where} \quad f(x) = \sum_n f_n x^n.$$

Although the notation of f_n is intuitive and useful when dealing with a single power series $f(x)$, it becomes a clutter when dealing with expressions of multiple power series including multiplication, exponentiation, and composition. Having 'the coefficient of x^n' as an operator rather than a variable name allows us to extract the nth coefficient of some expression of power series without having to state exactly the form of the nth coefficient. Where convenient to do so, we will state theorems in terms of the operator $[x^n]$ for consistency.

Remark 1. We will use the following properties without comment:

(1) $[x^n]x^k f(x) = [x^{n-k}]f(x)$,

(2) $\sum_{n\in\mathbb{Z}} x^n \cdot [z^n]f(z) = f(x)$. In our use case, we will always have $n < 0 \implies [x^n]f(x) = 0$.

(3) $[x^{n-1}]f'(x) = n[x^n]f(x)$, $n \geq 1$,

(4) $[x^n]$ is a linear operator.

Due to property (2), we formally need not specify the bounds of our sums as extra values will equate to 0. Regardless, we will specify the smallest possible bounds where possible.

2.3 Analysis

In analysis, every power series with a positive radius of convergence defines an analytic function within that radius. Thus we may consider a formal power series with a non-zero radius of convergence as an analytic function within its domain. This permits us to use the efficacious tools of complex analysis. As we will only consider power series centered at 0, we will state the coming results with that assumption baked in.

Given an analytic function $f(z) = \sum_{n=0}^{\infty} f_n z^n$, the nth coefficient is given by

$$[z^n]f(z) = \frac{f^{(n)}(0)}{n!}. \tag{2.4}$$

Omitting some conditions, we have the Cauchy integral formula and the Cauchy differentiation formula, applicable to an analytic $f(z)$ over a positively orientated boundary, denoted $\mathcal{O}^+$.

$$f(0) = \frac{1}{2\pi i}\oint_{\mathcal{O}^+} \frac{f(z)}{z}dz, \quad f^{(n)}(0) = \frac{n!}{2\pi i}\oint_{\mathcal{O}^+} \frac{f(z)}{z^{n+1}}dz. \tag{2.5}$$

Combining (2.5) and (2.4), we derive the Cauchy coefficient formula for analytic functions.

$$[z^n]f(z) = \frac{1}{2\pi i}\oint_{\mathcal{O}^+} \frac{f(z)}{z^{n+1}}dz. \tag{2.6}$$

Lastly, we state results of the residue theorem, again omitting some conditions.

$$\oint_{\mathcal{O}^+} f(z)dz = 2\pi i\,\mathrm{Res}(f), \quad \mathrm{Res}(f) = \frac{1}{(n-1)!}\lim_{z\to 0}\frac{d^{n-1}}{dz^{n-1}}(z^n f(z)). \tag{2.7}$$

These results will be used in Section 6.1 to prove the Lagrange inversion theorem.

3 The Lagrange Inversion Formula

3.1 Formal Statement

Theorem 3.1. *Given a formal power series $\phi(z) = \sum_{r=0}^{\infty} \phi_r z^r$, there exists a unique formal power series $y = y(x) = \sum_{n=0}^{\infty} y_n x^n$ satisfying*

$$y = x\phi(y), \tag{3.1}$$

where for any integer $n \geq 0$, we have

$$n[x^n]y = [z^{n-1}]\phi(z)^n, \tag{3.2}$$

for any integer $k > 0$, we have

$$n[x^n]y^k = k[z^{n-k}]\phi(z)^n, \tag{3.3}$$

and any for formal power series $H(z) = \sum_{r=0}^{\infty} h_r z^r$, we have

$$n[x^n]H(y) = [z^{n-1}]H'(z)\phi(z)^n. \tag{3.4}$$

It is clear that with $H(z) = z^k$, (3.4) implies (3.3), which with $k = 1$, implies (3.2). We will see in Section 6.1 that (3.4) is implied by (3.3).

3.2 Assumptions

In other literature, we often see the assumptions of $y_0 = 0$ and $\phi_0 \neq 0$, however, these conditions arise naturally and are formally redundant. $\phi_0 = 0$ implies $y_n = 0$ for all n, which corresponds to $y(x) = 0$. Thus (3.1) is trivially true as $\phi_0 = 0$ also implies $\phi(0) = 0$, giving us $0 = 0$. Since we multiply by x on the RHS, it is easy to see that there is no constant term. Thus by equating coefficients, there must also be no constant term on the LHS, i.e. $y_0 = 0$. We will see both of these cases in more detail in our proof in Section 6.1.

3.3 Equivalent Forms

The following forms are equivalent in that they are easily derivable from each other. We will show some but not all of these derivations.

If $f(z) = x$ is analytic at a with $f'(a) \neq 0$, then we have $y(x) = z$ given by

$$y(x) = a + \sum_{n=1}^{\infty} \frac{(x - f(a))^n}{n!} \lim_{z \to a} \frac{d^{n-1}}{dz^{n-1}} \left(\frac{z - a}{f(z) - f(a)} \right)^n \tag{3.5}$$

where $y(x)$ is the compositional inverse of $f(z)$.

Proof. We let $f(z) = z/\phi(z)$ with $\phi(0) \neq 0$ and take $a = 0$, noting $f(a) = f(0) = 0$. After evaluating the limit, we have

$$y(x) = \sum_{n=1}^{\infty} \frac{x^n}{n!} \lim_{z\to 0} \frac{d^{n-1}}{dz^{n-1}} \left(\frac{z}{z/\phi(z)}\right)^n = \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{1}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}} \phi(0)^n\right) x^n. \tag{3.6}$$

Lastly, applying (2.4) and extracting $[x^n]$ gives us (3.2)

$$y(x) = \sum_{n=1}^{\infty} \frac{1}{n} \left([z^{n-1}]\phi(z)^n\right) x^n, \qquad [x^n]y(x) = [z^{n-1}]\phi(z)^n. \tag{3.7}$$

□

In cases where $H'(z)$ is difficult to deal with, we have the following

$$[x^n]H(y) = [z^n]H(z)\phi(z)^{n-1} \left(\phi(z) - z\phi'(z)\right). \tag{3.8}$$

We will need (3.8) in the proof of the Vandermonde convolution generalisation in Section 5 as it is precisely a case with a difficult to deal with $H'(z)$.

We follow the proof of (3.8) by Gessel in [6, Section 2], modifying it to make the relationship explicit. We exploit the quotient rule and the fact that the coefficient of $[z^{-1}]$ in the derivative of any power series must be 0 to have our formula be in terms of $\phi'(z)$ instead of $H'(z)$.

Proof. Firstly, we notice that $y(x) = x\phi(y(x))$ can be rewritten as $x = y/\phi(y)$. We define $g(t) = t/\phi(t)$, where we have $g(y(x)) = x$. Thus, we have g being the compositional inverse of y with

$$g'(t) = \left(\frac{t}{\phi(t)}\right)' = \frac{\phi(t) - t\phi'(t)}{\phi(t)^2}. \tag{3.9}$$

Rewriting (3.4) in terms of g gives us

$$[x^n]H(y) = \frac{1}{n}[z^{n-1}]H'(z) \left(\frac{z}{g(z)}\right)^n = \frac{1}{n}[z^{-1}]\frac{H'(z)}{g(z)^n}. \tag{3.10}$$

By the quotient rule, we have

$$\left(\frac{H(z)}{g(z)^n}\right)' = \frac{H'(z)g(z)^n - ng(z)^{n-1}g'(z)H(z)}{g(z)^{2n}} = \frac{H'(z)}{g(z)^n} - \frac{nH(z)g'(z)}{g(z)^{n+1}}. \tag{3.11}$$

But, by properties of the derivative, the coefficient of $[z^{-1}]$ in the derivative of any power series must be 0. Thus after applying $[z^{-1}]$ to (3.11), we have

$$[x^n]H(y) = \frac{1}{n}[z^{-1}]\frac{H'(z)}{g(z)^n} = \frac{1}{n}[z^{-1}]\frac{nH(z)g'(z)}{g(z)^{n+1}}. \tag{3.12}$$

Rewriting $g(t)$ in terms of $\phi(t)$ and simplifying yields (3.8)

$$[x^n]H(y) = [z^{-1}]\frac{H(z)\phi(z)^{n+1}\left(\phi(z) - z\phi'(z)\right)}{z^{n+1}\phi(z)^2} = [z^n]H(z)\phi(z)^{n-1}(\phi(z) - z\phi'(z)). \tag{3.13}$$

□

For more exposition on equivalent forms and the derivations between them, we refer the reader to [6].

4 Examples

4.1 Lambert W Function

As previously mentioned, by defining $\phi(x) \coloneqq x/F(x)$, the Lagrange inversion allows us to find the compositional inverse of $F(x)$ in terms of another power series. As seen below, rearranging can be equivalent to this substitution.

Example 4.1 (Lambert W Function)**.** Given $f(y) = ye^y$, we solve for y. This will produce the *Lambert W Function*, the compositional inverse of $f(y)$, typically denoted $W(x)$. Let $x = ye^y$, rearranging into $y = xe^{-y}$ gives us the form of $y = x\phi(y)$ with $\phi(z) = e^{-z}$. After applying (3.2), we have:

$$[x^n]y = \frac{1}{n}[z^{n-1}]e^{-nz} = \frac{1}{n}[z^{n-1}]\sum_{k=0}^{\infty}\frac{(-n)^k}{k!}z^k = \frac{1}{n}\frac{(-n)^{n-1}}{(n-1)!} = \frac{(-n)^{n-1}}{n!}.$$

Now that we have the nth coefficients of $y(x)$, we simply sum over them.

$$W(x) = y = \sum_{n=1}^{\infty}\frac{(-n)^{n-1}}{n!}x^n.$$

The Lambert W function also has combinatorial interpretations. The *Cayley tree function* is given by $R(x) = -W(-x)$[8]. This function will make a brief appearance when we derive Abel's binomial theorem.

4.2 Generating Functions

A generating function is a way of presenting an infinite sequence of numbers as coefficients of a formal power series. Many combinatorial counting problems lend themselves to the Lagrange inversion formula as we can directly solve for the coefficients of their generating functions.

Example 4.2 (Catalan Numbers)**.** The *Catalan Numbers* are generated by solving

$$g = 1 + xg^2$$

for g. We will explore two substitutions that bring this into the form of $y = x\phi(y)$.

Remark 2. The nth *Catalan Number* is given by $C_n = \frac{1}{n+1}\binom{2n}{n}$.

- Let $y = g - 1$. This gives us $y = x(y+1)^2$, where $\phi(z) = (z+1)^2$. After applying (3.2) and the binomial theorem, we have

  $$[x^n]y = \frac{1}{n}[z^{n-1}](1+z)^{2n} = \frac{1}{n}[z^{n-1}]\sum_{k=0}^{2n}\binom{2n}{k}z^k.$$

  Extracting $[z^{n-1}]$ and rewriting gives us

  $$[x^n]y = \frac{1}{n}\binom{2n}{n-1} = \frac{1}{n}\frac{(2n)!}{(n-1)!(n+1)!} = \frac{1}{n+1}\frac{(2n)!}{(n!)^2} = \frac{1}{n+1}\binom{2n}{n} = C_n$$

Now we produce our sum and undo our substitution with $g = y + 1$.

$$y = \sum_{n=1}^{\infty} C_n x^n, \qquad g = \sum_{n=0}^{\infty} C_n x^n.$$

- As a proof of concept, we force a substitution of $y = xg$. Multiplying through by x gives us $xg = x + (xg)^2$, or $y = x + y^2$, which can be rewritten as

$$y = x(1-y)^{-1}.$$

with $\phi(z) = (1-z)^{-1}$. Applying (3.2) and expanding with the binomial theorem gives

$$[x^n]y = \frac{1}{n}[z^{n-1}](1-z)^{-n} = \frac{1}{n}[z^{n-1}]\sum_{k=0}^{\infty}\binom{-n}{k}(-1)^k z^k.$$

Extracting and rearranging yields

$$[x^n]y = \frac{1}{n}\binom{-n}{n-1}(-1)^{n-1} = \frac{1}{n}\binom{2n-2}{n-1}.$$

Lastly, we construct our sum and undo the substitution, granting us the same result.

$$xg = \sum_{n=1}^{\infty}\frac{1}{n}\binom{2n-2}{n-1}x^n = \sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}x^{n+1} = \sum_{n=0}^{\infty} C_n x^{n+1},$$

$$g = \sum_{n=0}^{\infty} C_n x^n.$$

Example 4.3 (Motzkin Numbers)**.** The *Motzkin Numbers* are generated by solving

$$h = 1 + xh + x^2h^2$$

for h. This can be done by multiplying through by x and the substitution $y = xh$ to get our desired form.

Remark 3. The nth *Motzkin Number* is given by $M_n = \sum_{k=0}^{\lfloor n/2 \rfloor}\binom{n}{2k}C_k$.

With $y = xh$, we have:

$$h = 1 + xh + x^2h^2, \quad xh = x(1 + xh + x^2h^2), \quad y = x(1 + y + y^2)$$

where $\phi(z) = (1 + z + z^2)$. Let us first compute $\phi(z)^n$. Repeated application of the binomial theorem gives

$$\phi(z)^n = (1+z+z^2)^n = \sum_{k=0}^{n}\binom{n}{k}z^k(1+z)^k = \sum_{k=0}^{n}\binom{n}{k}z^k\sum_{l=0}^{k}\binom{k}{l}z^l = \sum_{k=0}^{n}\sum_{l=0}^{k}\binom{n}{k}\binom{k}{l}z^{k+l}.$$

Letting $m = k + l$ allows us to collect powers of z, with $0 \le m \le 2n$. We use $k = m - l$ to simplify the inner sum. $\binom{m-l}{l}$ bounds l above by $\lfloor m/2 \rfloor$. These bounds are formally unnecessary as excess terms are simply 0.

$$\phi(z)^n = \sum_{m=0}^{2n} z^m \sum_{\substack{k+l=m \\ k,l \ge 0}}\binom{n}{k}\binom{k}{l} = \sum_{m=0}^{2n} z^m \sum_{l=0}^{\lfloor m/2 \rfloor}\binom{n}{m-l}\binom{m-l}{l}.$$

Now with (3.2), we have

$$
\begin{aligned}
[x^n]y &= \frac{1}{n}[z^{n-1}]\phi(z)^n \\
&= \frac{1}{n}[z^{n-1}]\sum_{m=0}^{2n} z^m \sum_{k=0}^{\lfloor m/2\rfloor} \binom{n}{m-k}\binom{m-k}{k} = \frac{1}{n}\sum_{k=0}^{\lfloor \frac{n-1}{2}\rfloor} \binom{n}{n-1-k}\binom{n-1-k}{k}.
\end{aligned}
$$

This can be rewritten in terms of the Catalan numbers by multiplying by $\frac{(2k)!}{(2k)!}$ and rewriting.

$$
\begin{aligned}
[x^n]y &= \sum_{k=0}^{\lfloor \frac{n-1}{2}\rfloor} \frac{(n-1)!}{k!(k+1)!(n-1-2k)!}\cdot\frac{(2k)!}{(2k)!} \\
&= \sum_{k=0}^{\lfloor \frac{n-1}{2}\rfloor} \frac{(n-1)!}{(2k)!(n-1-2k)!}\cdot\frac{1}{k+1}\frac{(2k)!}{(k!)^2} = \sum_{k=0}^{\lfloor \frac{n-1}{2}\rfloor}\binom{n-1}{2k}C_k = M_{n-1}.
\end{aligned}
$$

Lastly, we produce our series and undo our substitution.

$$
y = \sum_{n=1}^{\infty} x^n M_{n-1}, \qquad g = \sum_{n=0}^{\infty} x^n M_n.
$$

4.3 Solubility of the Quintic

Although the quintic (and higher order polynomials) are insoluble by radicals, [3] we show that permitting infinite sums changes this, as we merely need to evaluate the polynomial's inverse at 0. For our particular example, this is simply a direct extension of what we did for the Catalan numbers. We will use the same variables for consistency.

Example 4.4 (Quintic)**.** Suppose we want to solve the polynomial equation $g^5 - g + x = 0$ for g. This can be rearranged into $g = x(1-g^4)^{-1}$. With $\phi(z) = (1-z^4)^{-1}$, we apply (3.2) and expand with the binomial theorem yielding

$$
[x^n]g = \frac{1}{n}[z^{n-1}](1-z^4)^{-n} = \frac{1}{n}[z^{n-1}]\sum_{k=0}^{\infty}\binom{-n}{k}(-1)^k z^{4k} = \frac{1}{n}[z^{n-1}]\sum_{k=0}^{\infty}\binom{n+k-1}{k}z^{4k}.
$$

We notice that we have an exponent of $4k$, i.e. every $[z^{n-1}]$ not of this form will be 0. So we let $n = 4N+1$, this gives us

$$
[x^{4N+1}]g = \frac{1}{4N}[z^{4N}]\sum_{k=0}\binom{4N+k-1}{k}z^{4k}.
$$

Extracting our coefficient with $N = k$ and constructing our sum gives us

$$
[x^{4N+1}]g = \frac{1}{4N}\binom{5N}{N}, \quad g(x) = \sum_{n=0}^{\infty}\frac{1}{4n}\binom{5n}{n}x^{4n+1}.
$$

i.e. for a given constant x, this sum solves our polynomial. It can be shown via convergence tests that we have a radius of convergence of $|x| \le 4\cdot 5^{-5/4} \approx 0.53$.

More generally, we leave to the reader to verify that even for $p > 5$, this can be done for any polynomial equation $g^p - g + x = 0$ with g given by

$$
g = \sum_{n=0}^{\infty}\frac{1}{(p-1)n}\binom{pn}{n}x^{(p-1)n+1}, \quad |x| \le (p-1)p^{-p/(p-1)}.
$$

4.4 Combinatorial Identities

We can apply the binomial theorem and Lagrange inversion theorem in creative ways to produce combinatorial identities. A surprisingly effective approach is to trivially restate some expression in terms of its components. Expanding component wise and equating to the expansion of the original expression can yield a non-trivial identity as we see below.

Remark 4. Let $U(x) = \sum_{p\geq 0} u_p x^p$ and $V(x) = \sum_{q\geq 0} v_q x^q$ be formal power series. Then

$$[x^n]U(x)V(x) = \sum_{p+q=n} [x^p]U(x) \cdot [x^q]V(x).$$

This is a direct consequence of power series multiplication.

Example 4.5 (Abel's Binomial Theorem)**.** By computing in two different ways the coefficient

$$[x^n]e^{(\alpha+\beta)y} = [x^n]e^{\alpha y} \cdot e^{\beta y},$$

where $y = xe^y$ is the Cayley tree function, one derives the following collection of identities

$$(\alpha+\beta)(n+\alpha+\beta)^{n-1} = \alpha\beta \sum_{k=0}^{n} \binom{n}{k} (\alpha+\beta)^{k-1}(\beta+n-k)^{n-k-1}.$$

[4]

Proof. We observe that $y = xe^y$ is of the form $y = x\phi(y)$ with $\phi(z) = e^z$, applying (3.4) with $H(z) = e^{(\alpha+\beta)x}$, $H'(z) = (\alpha+\beta)e^{(\alpha+\beta)x}$, we have the following:

For the LHS, we let $\lambda = \alpha + \beta$.

$$[x^n]e^{\lambda y} = \frac{\lambda}{n}[z^{n-1}]e^{(\lambda+n)z} = \frac{\lambda}{n}[z^{n-1}]\sum_{k=0}^{\infty}\frac{(\lambda+n)^k}{k!}z^k = \frac{\lambda(\lambda+n)^{(n-1)}}{n!} = \frac{(\alpha+\beta)(\alpha+\beta+n)^{(n-1)}}{n!}.$$

For the RHS, we consider $[x^n]e^{(\alpha+\beta)y} = [x^n]e^{\alpha y}e^{\beta y}$ and apply remark 4.

$$[x^n]e^{\alpha y}e^{\beta y} = \sum_{\substack{n=k+l \\ k,l\geq 0}} [x^k]e^{\alpha y}[x^l]e^{\beta y} = \sum_{\substack{n=k+l \\ k,l\geq 0}} \frac{\alpha(\alpha+k)^{k-1}}{k!}\frac{\beta(\beta+l)^{l-1}}{l!}.$$

Substituting $l = n - k$ and multiplying by $\frac{n!}{n!}$ allows this to be rewritten as

$$[x^n]e^{\alpha y}e^{\beta y} = \frac{\alpha\beta}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}(\alpha+k)^{k-1}(\beta+n-k)^{n-k-1} = \frac{\alpha\beta}{n!}\sum_{k=0}^{n}\binom{n}{k}(\alpha+k)^{k-1}(\beta+n-k)^{n-k-1}.$$

Cancelling $\frac{1}{n!}$ on both sides yields

$$(\alpha+\beta)(n+\alpha+\beta)^{n-1} = \alpha\beta \sum_{k=0}^{n} \binom{n}{k} (\alpha+\beta)^{k-1}(\beta+n-k)^{n-k-1}.$$

□

5 A Generalisation of Vandermonde's Convolution

We study a variation of the identity found in [7, equation 11]. We present a novel approach and provide a more elegant proof via Lagrange inversion.

Example 5.1 (A Generalisation of Vandermonde's Convolution)**.**

$$\sum_{n=0}^{\infty}\binom{m+n(\lambda+1)}{n}z^n = \left(\sum_{n=0}^{\infty}\binom{n(\lambda+1)}{n}\frac{z^n}{1+\lambda n}\right)^m\left(\sum_{n=0}^{\infty}\binom{m+n(\lambda+1)}{n}z^n\right) \tag{5.1}$$

In the coming proof, we exploit the trivial fact that

$$\frac{(1+y)^{m+1}}{1-\lambda y} = (1+y)^m\cdot\frac{1+y}{1-\lambda y}.$$

We then use the Lagrange inversion formula to compute the LHS and RHS via different means, deriving a non-trivial identity that is laborious to prove otherwise [7].

Proof. We let $z = y/(1+y)^{\lambda+1}$ and rearrange so that $z = y/(1+y)^{\lambda+1}$. Applying (3.2) with $\phi(z) = (1+z)^{\lambda+1}$ and expanding with the binomial theorem gives us

$$[x^n]y = \frac{1}{n}[z^{n-1}](1+z)^{n(\lambda+1)} = \frac{1}{n}[z^{n-1}]\sum_{k=0}^{\infty}\binom{n(\lambda+1)}{k}z^k.$$

After extracting $[z^{n-1}]$ and rewriting, we have

$$[x^n]y = \frac{1}{n}\binom{n(\lambda+1)}{n-1} = \binom{n(\lambda+1)}{n}\frac{1}{1+n\lambda}$$

allowing us to produce our first sum

$$y = \sum_{n=1}^{\infty}\binom{n(\lambda+1)}{n}\frac{z^n}{1+\lambda n}, \qquad 1+y = \sum_{n=0}^{\infty}\binom{n(\lambda+1)}{n}\frac{z^n}{1+\lambda n}. \tag{5.2}$$

Next, we apply (3.8) with $H(z) = \dfrac{(1+z)^{m+1}}{1-\lambda z}$, first noting

$$\phi(z) - z\phi'(z) = (1+z)^{\lambda+1} - z(\lambda+1)(1+z)^{\lambda} = (1+z)^{\lambda}(1-z\lambda). \tag{5.3}$$

Now with the above, we have

$$[x^n]H(y) = [z^n]\frac{(1+z)^{m+1}}{1-\lambda z}(1+z)^{(n-1)(\lambda+1)+\lambda}(1-z\lambda) = (1+z)^{m+n(\lambda+1)}.$$

Applying the binomial theorem and extracting our coefficient gives us

$$[x^n]\frac{(1+y)^{m+1}}{1-\lambda y} = [z^n]\sum_{k=0}^{\infty}\binom{m+n(\lambda+1)}{k}z^k = \binom{m+n(\lambda+1)}{n}$$

allowing us to produce our second sum

$$\frac{(1+y)^{m+1}}{1-\lambda y} = \sum_{n=0}^{\infty}\binom{m+n(\lambda+1)}{n}z^n. \tag{5.4}$$

We observe that
$$\frac{(1+y)^{m+1}}{1-\lambda y} = (1+y)^m \cdot \frac{1+y}{1-\lambda y}.$$
Lastly, to the LHS we apply (5.4) and to the RHS we apply (5.2) as well as (5.4) with $m = 1$, yielding (5.1)
$$\sum_{n=0}^{\infty} \binom{m+n(\lambda+1)}{n} z^n = \left(\sum_{n=0}^{\infty} \binom{n(\lambda+1)}{n} \frac{z^n}{1+\lambda n} \right)^m \left(\sum_{n=0}^{\infty} \binom{m+n(\lambda+1)}{n} z^n \right).$$
□

Remark 5. We emphasise that (3.8) is crucial here as a direct application of (3.4) does not seem to work. We leave this as a challenge for the keen reader.

6 Proof

We will prove Theorem 3.1 by firstly showing that a unique solution exists for equation (3.1). We see that by expanding and equating coefficients of each power series, one can explicitly compute each y_n one by one. This begs the question: why do we need Lagrange inversion if these coefficients can be computed explicitly? Working with this system of equations is a hassle. Lagrange inversion shortcuts this process and directly provides a formula for y_n without having to solve a system of equations. We state this system of equations purely to verify existence and uniqueness.

We then argue that the conditions of (2.6) apply and using a change of variables, produce our result. "In the context of complex analysis, [Lagrange inversion] appears as nothing but an avatar of the change-of-variable formula."[4] The Lagrange inversion theorem can also be proved via residues, induction and by many other means[6, Section 4].

6.1 Proof via Cauchy's Coefficient Formula

Proof. We first show $y = x\phi(y)$ yields a unique solution with $y(x) = \sum_{n=0}^{\infty} y_n x^n, \phi(z) = \sum_{n=0}^{\infty} \phi_n z^n$.
$$\sum_{n=0}^{\infty} y_n x^n = x \sum_{n=0}^{\infty} \phi_n \left(\sum_{k=0}^{\infty} y_k x^k \right)^n$$
We now expand the RHS, noting that the RHS has no constant term, thus the LHS has no constant term. Since we have $y_0 = 0$, we omit y_0 from the expansion.
$$\begin{aligned}
&n = 0: \ x\phi_0 \\
&n = 1: \ x\phi_1(y_1x + y_2x^2 + y_3x^3 + \cdots) &&= \phi_1y_1x^2 + \phi_1y_2x^3 + \phi_1y_3x^4 + ... \\
&n = 2: \ x\phi_2(y_1x + y_2x^2 + \cdots)^2 &&= \phi_2y_1^2x^3 + 2\phi_2y_1y_2x^4 + ... \\
&n = 3: \ x\phi_3(y_1x + \cdots)^3 &&= \phi_3y_1^3x^4 + ... \\
&\quad \vdots &&\quad \vdots \\
&n = k: \ x\phi_k(y_1x + \cdots)^k &&= \phi_ky_1^kx^{k+1} + ...
\end{aligned}$$

Equating coefficients and solving for y_n yields

$$\begin{aligned} y_0 &= 0 \\ y_1 &= \phi_0 \\ y_2 = \phi_1 y_1 &= \phi_0\phi_1 \\ y_3 = \phi_1 y_2 + \phi_2 y_1^2 &= \phi_0\phi_1^2 + \phi_0^2\phi_2 \\ y_4 = \phi_1 y_3 + 2\phi_2 y_1 y_2 + \phi_3 y_1^3 &= \phi_0\phi_1^3 + 3\phi_0^2\phi_1\phi_2 + \phi_0^3\phi_3 \\ &\vdots \end{aligned}$$

Let us consider the case where $\phi_0 = 0 = y_1$. We show via strong induction that this implies that for all n, we have $y_n = 0$. We notice that since $y_0 = 0$, a given y_m depends polynomially on $y_{m-1}, ..., y_1$ with no constant term. So y_m satisfies that condition that $y_1 = 0, ..., y_{m-1} = 0$ implies $y_m = 0$. Thus with our base case of $y_1 = 0$, we have $y_n = 0$ for all n which implies $y(x) = 0$. We now resume with $\phi_0 \neq 0$.

Each y_m depends on the coefficients of only the first $m-1$ terms of $\phi(x)$, thus for a given m, we need only consider $\phi_{[m]} := \sum_{n=0}^{m-1} \phi_n x^n$ to calculate y_m, which is a polynomial, i.e. analytic. Consider

$$y = x\phi_{[m]}(y), \quad x = \frac{y}{\phi_{[m]}(y)} := f(y)$$

Since $\phi_0 \neq 0$ we know $\phi_{[m]}(0) \neq 0$. Thus by the analyticity of $\phi_{[m]}(z), f(y)$ is also analytic about 0. Lastly, by [5, Theorem 7.5 - Inverse Mapping Theorem], we have the existence and analyticity of the inverse, $g(x) = y = g(f(y))$. As this holds for all m, for any formal power series $\phi(z)$ (which may not be analytic), we can always consider the truncated series $\phi_{[m]}(z)$(which is always analytic), thus without loss of generality, we can assume that $\phi(x)$ is a polynomial. With this in consideration, we are permitted to use (2.6). We emphasize that since we do not require $\phi(z)$ to be analytic, $\phi(z)$ is permitted to be any formal power series as we only use the analyticity of $\phi_{[m]}(z)$.

To prove (3.2), we apply (2.6) on $[z^{n-1}]y(z)' = n[z^n]y(z)$. After applying a change of variables $z = y/\phi(y)$ with $y'dz = dy$ we have

$$n[z^n]y = \frac{1}{2\pi i}\oint_{\mathcal{O}+} \frac{y'}{z^n}dz = \frac{1}{2\pi i}\oint_{\mathcal{O}+} \frac{\phi(y)^n}{y^n}dy \tag{6.1}$$

We then apply (2.7) and (2.4) yielding

$$n[z^n]y = \mathrm{Res}\left(\frac{\phi(y)^n}{y^n}\right) = \frac{1}{(n-1)!}\lim_{y\to 0}\frac{d^{n-1}}{dy^{n-1}}\frac{y^n\phi(y)^n}{y^n} = \frac{1}{(n-1)!}\frac{d^{n-1}}{dy^{n-1}}\phi(0)^n = [y^{n-1}]\phi(y)^n.$$

To prove (3.3), we do the same with $[z^{n-1}]\left(y(z)^k\right)' = n[z^n]y(z)^k$, noting that $(y^k)' = ky^{k-1}y'$. We again let $z = y/\phi(y)$ with $dy = y'dz$.

$$n[z^n]y^k = \frac{1}{2\pi i}\oint_{\mathcal{O}+} \frac{(y^k)'}{z^n}dz = \frac{1}{2\pi i}\oint_{\mathcal{O}+} \frac{ky^{k-1}y'}{z^n}dz = \frac{k}{2\pi i}\oint_{\mathcal{O}+} y^{k-n-1}\phi(y)^n dy. \tag{6.2}$$

Applying (2.7) and (2.4) gives us

$$n[z^n]y^k = \mathrm{Res}\left(\frac{\phi(y)^n}{y^{n-k-1}}\right) = \frac{k}{(n-k)!}\lim_{y\to 0}\frac{d^{n-k}}{dy^{n-k}}y^{n-k-1}\frac{\phi(y)^n}{y^{n-k-1}} = \frac{k}{n-k!}\frac{d^{n-k}}{dy^{n-k}}\phi(0)^n = k[y^{n-k}]\phi(y)^n$$

By linearity of the coefficient operator, (3.4) is implied by (3.3) as shown below [9]

$$n[x^n]H(y) = \sum_{k=0}^{\infty} h_k n[x^n]y^k = \sum_{k=0}^{\infty} h_k k[z^{n-k}]\phi(z)^n = \sum_{k=0}^{\infty} h_k [z^{n-1}]kz^{k-1}\phi(z)^n. \tag{6.3}$$

After factoring a rewriting in terms of derivatives, we have

$$n[x^n]H(y) = [z^{n-1}]\left(\sum_{k=0}^{\infty} h_k (z^k)'\right)\phi(z)^n = [z^{n-1}]H'(z)\phi(z)^n. \tag{6.4}$$

□

As mentioned in Section 3.3, there are many equivalent forms which imply one another [6].

7 Afterword

The Lagrange inversion theorem stands as a versatile and reliable tool in providing elegant solutions to a variety of problems in combinatorics and analysis. We have shown its ability to easily extract coefficients from generating functions and produce non-trivial combinatorial identities. In particular, we presented a novel proof of a generalisation of Vandermonde's convolution (see [7] for original proof) through direct application of the Lagrange inversion formula. Even in the case where the standard Lagrange or Bürmann forms fail to follow through, we have shown that we can often use equivalent forms to avoid common issues. We were also able to provide an explicit solution to an infinite family of polynomials commonly known to be insoluble by radicals [3]. Although seemingly restrictive, we find that the condition of $y = x\phi(y)$ either appears naturally or can easily be forced, as we have seen in the application of inverting any power series.

References

[1] Manuel Bronstein, Robert M. Corless, James H. Davenport, and D. J. Jeffrey. Algebraic properties of the Lambert W function from a result of Rosenlicht and of Liouville. *Integral Transforms and Special Functions*, 19(10):709–712, 2008.

[2] Timothy Y. Chow. What is a closed-form number?, 1998.

[3] Harold M. Edwards. *Galois Theory*. Springer New York, NY, 1984.

[4] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[5] Klaus Fritzsche and Hans Grauert. *From Holomorphic Functions to Complex Manifolds*. Springer New York, NY, 2002.

[6] Ira M. Gessel. Lagrange Inversion. *Journal of Combinatorial Theory, Series A*, 144:212–249, 2016. Fifty Years of the Journal of Combinatorial Theory.

[7] H. W. Gould. Some Generalizations of Vandermonde's Convolution. *The American Mathematical Monthly*, 63(2):84–91, 1956.

[8] nLab authors. Lambert W-function. `https://ncatlab.org/nlab/show/Lambert+W-function`, May 2024. Revision 2.

[9] Erlang Surya and Lutz Warnke. Lagrange Inversion Formula by Induction. *The American Mathematical Monthly*, 130(10):944–948, 2023.